

畳み込みニューラルネットワークを用いた微小眼球運動の検出

Convolutional Neural Network for Microsaccade Detection

惠本序珠亜*¹
Emoto, Joshua

平田 豊*^{1,2}
Hirata, Yutaka

*¹ 中部大学大学院 工学部 情報工学専攻
Department of Computer Science, Chubu University Graduate School of Engineering

*² 中部大学 工学部 ロボット理工学科
Department of Robotic Science and Technology, Chubu University College of Engineering

Microsaccade (MSC) is a miniature rapid eye movement, which leads to a variety of potential applications as those related to visual attention and brain diseases in human. However, accurate detection of MSCs is difficult because of its very small amplitude and short duration. In this study, we propose a novel MSC detection algorithm employing a convolutional neural network. We demonstrate that the proposed algorithm enables to detect MSCs from noisy eye position data in real time with high accuracy and robustness.

1. はじめに

近年、マイクロサッカド(MSC)と呼ばれる微小で高速な眼球運動が様々な分野で注目されている。すなわち、MSCには人の空間的注意が反映されると考えられており、心理学、神経科学分野のみならずマーケティング、自動車安全技術分野などでの応用が期待されている。眼球運動データからMSCを抽出するために、すでにいくつかの手法が提案されている [Engbert 2003, Bettenbühl 2009, Otero-Millan 2014]。しかしながら、これらの手法は頭部が固定された状態などのノイズ混入の比較的小さいデータからは安定してMSCを検出することができるものの、ノイズの多いデータやノイズの特性変動が激しいデータからMSCを検出することは想定されておらず、より実用的な場面で計測された眼球運動データからは誤検出や検出漏れが発生しやすくなっている。MSCの応用可能性が広がる中、ノイズの激しいデータからも頑健にMSCを検出できるアルゴリズムが求められている。

最近、ディープラーニングという機械学習の手法が画像処理や音声処理における特徴検出や認識において高い性能を発揮することが示されている。特に、畳み込みニューラルネットワーク(ConvNet)は画像中の抽象的な特徴をうまくとらえるものとして広く応用され、顕著な成功をおさめている。そこで本研究では、ConvNetを眼球運動データ内の特定波形検出用にデザインし、ノイズの激しいデータからも安定してMSCを検出する手法を提案する。

2. 従来手法

現在、MSCを簡便に検出する方法として広く利用されている手法は[Engbert 2003]により提案されたものである。この手法は眼球位置波形より算出した速度波形から、ノイズレベルに合わせて調節される閾値によりMSCを検出する。最近では、眼球速

度波形と加速度波形から算出した特徴量からクラスタリングによりMSCを検出する手法が提案されている[Otero-Millan 2014]。これらの手法では、速度波形や加速度波形によりMSCの特徴を捉えているが、これらから得られる特徴量だけではMSCはうまく検出できても、MSCでないノイズ成分も検出してしまい、MSCとノイズ成分を区別できず、誤検出を抑えることが難しいと考えられる。ノイズの激しいデータからMSCを安定して検出するためには、MSCの特徴と同時にMSCでない成分の特徴もとらえる手法が有効である。

3. 提案手法

本手法では、ConvNetを応用し、眼球位置波形を入力としてMSCの特徴と同時に、MSCでない成分の特徴を学習することで、ノイズに頑健なMSC検出を実現することを考える。現在、物体検出の手法として主流となっているのはR-CNNという手法である[中山 2015]。そこで、R-CNNから派生した物体を高速かつ高精度で検出することを目的とした手法であるFaster R-CNN[Shaoqing 2015]を参考にConvNetの構造と学習方法を決定した。

3.1 検出器の構造

Faster R-CNNでは、画像から物体領域を抽出するニューラルネットワーク構造としてRegion Proposal Networkが提案されている。そこで、本手法で用いるConvNetの構造はRegion Proposal Networkを参考にした。

本手法で用いたConvNetは特徴抽出部とMSC識別部の二つに分けることができる。

特徴抽出部は2層の畳み込み層で構成されている。2層の畳み込み層により、特徴抽出部は、従来手法で用いられていた眼球速度と加速度成分を入力波形から抽出するフィルターを表現することができるとともに、データセットを基に最適なフィルターを学習することができると考えられる。図1に特徴抽出部の構造を示す。図の上部が特徴抽出部を表し、下部が入力および出力波形を表す。図中の直方体は波形が複数あることを表し、その右下の数字が波形の数を表す。カーネル数やその点数、活性化関数についても図中に示した。

連絡先:

惠本序珠亜, 中部大学大学院工学部 情報工学専攻,
emoto.joshua@gmail.com

平田 豊, 中部大学工学部 ロボット理工学科,
yutaka@isc.chubu.ac.jp

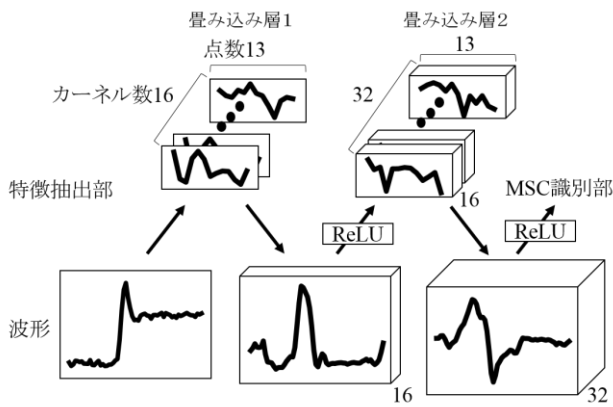


図1 特徴抽出部

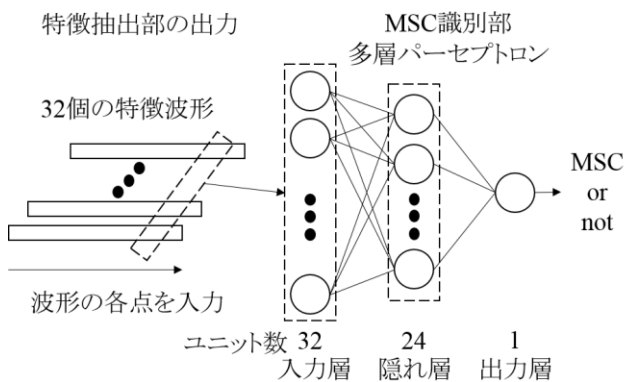


図2 MSC識別部

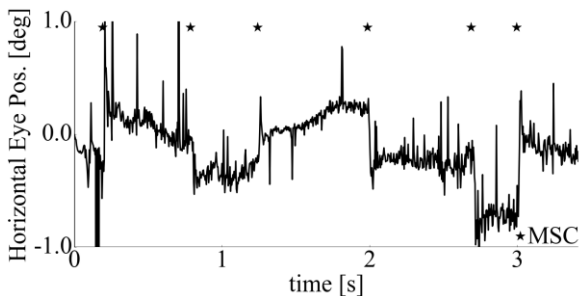


図3 MSCを含む眼球運動データ例

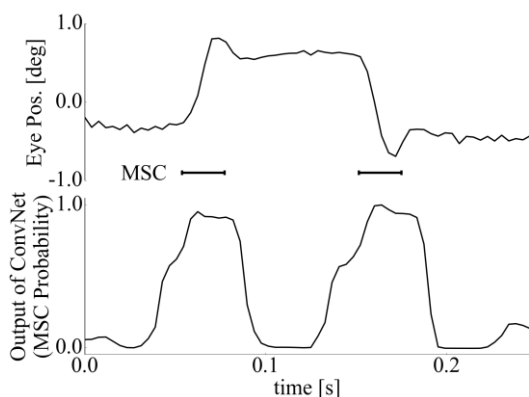


図4 ConvNetの出力例

MSC 識別部では、特徴抽出部から出力されたデータの各点が MSC であるか否かを識別する。図 2 に MSC 識別部の構造

を示す。MSC 識別部は特徴抽出部から出力された 32 個の特徴波形の各点を特徴量として多層パーセプトロンに入力する。よって、波形の長さが N 点であれば、同じ多層パーセプトロンで N 回の識別を行う。

本手法で用いる ConvNet はすべて畳み込み処理により構成されている。そのため、入力データの点数を固定する必要がなく、任意の点数の波形を入力することができ、波形の各点について MSC の真偽識別をすることができる。

3.2 学習方法

本手法の ConvNet は入力された波形の各点で識別結果を出力する。そのため、各入力点における MSC 識別の誤差をすべて用いて ConvNet を学習させると、波形に多く含まれる MSC でない成分の学習に偏ってしまう。そこで、全出力のうち、MSC とそうでない成分の割合がおよそ一対一となるように出力をサンプリングする。そして、サンプリングした出力の誤差についてのみ誤差逆伝搬法により ConvNet の学習を行う。

4. 結果

注視状態の眼球運動を計測して ConvNet の学習に必要なデータセットを作成した。図 3 に計測した眼球運動データ(眼球位置時系列)の例を示す。データセットの教師信号であるラベルは MSC であるデータ点を目視で確認して作成した。データは 2 名の被験者から取得し、眼球位置時系列の垂直・水平成分のうち、MSC が多く含まれる水平方向の眼球位置波形のみを学習に使用した。ConvNet を学習させる際の入力波形については固定長とし、平均が 0、分散が 1 となるように正規化した。図 4 に ConvNet の出力例を示す。ConvNet を学習させた結果、MSC の検出精度は 94.3% となった(5314MSC 中 5010 個)。

5. 結論

眼球運動データ内の特定波形検出用に ConvNet をデザインし、MSC を検出する手法を提案した。また、実データを用いて提案手法を評価した。提案手法は眼球運動データから最適なフィルターを学習し、MSC を検出するために有効な特徴量を抽出することで、ノイズの激しいデータからも頑健な MSC 検出を可能とすることが実証された。

参考文献

[Engbert 2003] Engbert R. Kliegl R.: Microsaccades uncover the orientation of covert attention, *Vision Research*, Apr;43(9):1035-45, 2003.

[Bettenbühl 2009] Bettenbühl M. Paladini C. Mergenthaler K. Kliegl R. Engbert R. Holschneider M.: Microsaccade characterization using the continuous wavelet transform and principal component analysis, *J. Eye Movement Res.*, 3(5):1, 1-14, 2009.

[Otero-Millan 2014] Jorge Otero-Millan, Jose L. Alba Castro, Stephen L. Macknik, Susana Martinez-Conde: Unsupervised clustering method to detect microsaccades, *J. Vision*, Vol.14, 18, 2014.

[中山 2015] 中山英樹: 深層畳み込みニューラルネットワークによる画像特徴抽出と転移学習, 電子情報通信学会音声研究会 7 月研究会, 2015.

[Shaoqing 2015] Shaoqing Ren, Kaiming He, Ross Girshick, Jian Sun: Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks, *Advances in Neural Information Processing Systems*, 28:91-99, 2015.